A CLASS OF NONLINEAR OPTIMIZATION PROBLEMS WITH POTENTIALS

JIAKUN LIU

ABSTRACT. In this paper, we introduce a class of nonlinear optimization problems. Under mild assumptions, we obtain the existence of potential functions and show that the potential function is a generalized solution of a Monge-Ampère type equation. We also present some interesting applications in optimal transportation and geometric optics problems.

1. INTRODUCTION

In a previous paper [16], we proved that the reflector problem in the near field case is a nonlinear optimization problem. In this paper we introduce a class of nonlinear optimization problems with potentials and prove the existence of optimal mappings. These problems are natural extensions of the optimal transportation problem, which is a class of linear optimization problems extensively studied in recent years. In particular, we give more examples of nonlinear optimization problems arising in reflector and refractor problems. Our results, together with [12, 25], show that the light reflection and refraction processes can be formulated to nonlinear optimization problems.

Let's consider a class of optimization problems of maximizing the following functional

$$(1.1) \qquad I(u,v) := \int_{U \times V} F(x,y,u(x),v(y)) d\gamma,$$

where U, V are two bounded domains in $\mathbb{R}^n$, F is a function in $\mathbb{R}^{2n+2}$, $d\gamma$ is a measure on $U \times V$, which has dx, dy as its marginals, and (u,v) is a pair of functions contained in the constraint set

$$(1.2) \qquad K = \{(u,v) \in C(U) \times C(V) \, : \, \phi(x,y,u,v) \le 0 \text{ in } U \times V\},$$

where ϕ is the nonlinear constraint function. More generally, we may also consider U, V subsets of a Riemannian manifold $(\mathcal{M}, g)$. Some interesting examples and applications of the problem (1.1)–(1.2) are included in Section 5.

Write $F = F(x,y,t,s)$ and $\phi = \phi(x,y,t,s)$, where x,y,t,s are independent variables. Use the subscripts to denote the partial derivatives, i.e., $F_{x_i} = \partial F/\partial x_i$, $\phi_t = \partial \phi/\partial t$, etc.

We always assume the following hypotheses: The function F is C^1 smooth and monotone increasing in t, s, namely

$$(1.3) \qquad\qquad F_t \geq 0, \quad F_s \geq 0 \quad \text{in } U \times V \times \mathbb{R}^2,$$

and the constraint $\phi = \phi(x, y, t, s)$ is C^1 smooth and strictly increasing in t, s. Note that by the implicit function theorem, the constraint $\phi(x, y, u, v) = 0$ can always be written as

$$\phi\left(x, y, -\varphi(x, y, v), v\right) = 0,$$

for some function $\varphi = \varphi(x, y, s)$, which is C^1 smooth in s. Since ϕ is strictly increasing in t, without loss of generality, the constraint $\phi \leq 0$ can be written as

$$(1.4) \qquad\qquad \phi(x, y, u, v) = u + \varphi(x, y, v) \leq 0.$$

Correspondingly, the function $\varphi = \varphi(x, y, s)$ is strictly increasing in s. We assume further that there exists a constant $\theta_0 > 0$ such that

$$(1.5) \qquad\qquad \varphi_s \geq \theta_0 \quad \text{in } U \times V \times \mathbb{R}.$$

In order to state our main results, we introduce the following conditions and define some terminology.

(H1) For each $x_0 \in U$, for any $(p, t) \in \mathbb{R}^n \times \mathbb{R}$, there is at most one pair $(y, s) \in \mathbb{R}^n \times \mathbb{R}$ such that

$$(\varphi_x, \varphi)(x_0, y, s) = -(p, t),$$

namely, $\varphi_x(x_0, y, s) = -p$ and $\varphi(x_0, y, s) = -t$.

(H2) For any $(x, y, s) \in U \times V \times \mathbb{R}$,

$$\det\left(\varphi_{xy} - \frac{1}{\varphi_s}\varphi_{xs} \otimes \varphi_y\right) \neq 0.$$

Definition 1.1. *A pair $(u, v) \in K$ is called a* dual pair *with respect to ϕ if*

$$(1.6) \qquad\qquad u(x) = \sup\{t \,:\, \phi(x, y, t, v(y)) \leq 0, \quad \forall y \in V\},$$
$$v(y) = \sup\{s \,:\, \phi(x, y, u(x), s) \leq 0, \quad \forall x \in U\}.$$

If furthermore $I(u, v) = \sup_K I$, (u, v) is called a dual maximizing pair *of I, and the functions u, v are called* potential functions *in the nonlinear optimization (1.1)–(1.2).*

Definition 1.2. *Let $(u, v) \in K$ be a dual maximizing pair of I. A mapping $T : U \to V$ is called an* optimal mapping *associated to (u, v) if T solves the equation*

$$(1.7) \qquad\qquad \phi(x, Tx, u(x), v(Tx)) = 0$$

almost everywhere on U.

The following theorem gives the existence of potential functions and optimal mappings in the nonlinear optimization problem (1.1)–(1.2). Their regularity properties will be investigated in a subsequent paper [17].

Theorem 1.1. *In addition to hypotheses (1.3)–(1.5) and (H1)–(H2), we assume the following mixed condition on F and φ: for any $(u, v) \in K$, there holds*

$$(1.8) \qquad \int_{U \times V} \left\{ -F_t(x, y, u(x), v(y)) + \frac{F_s}{\varphi_s}(x, y, u(x), v(y)) \right\} d\gamma = 0.$$

Then there exists a dual maximizing pair $(u, v) \in K$ of I, and a unique optimal mapping T associated to (u, v) uniquely determined almost everywhere on U.

Let us consider a special case, which has many practical applications, that

$$(1.9) \qquad F(x, y, u, v) = u(x)f(x) + \varphi(x, y, v)g(y),$$

where $f \in L^1(U)$, $g \in L^1(V)$, and φ is the constraint in (1.4). By the monotonicity (1.3)–(1.5), we assume that f, g are positive functions. The condition (1.8) implies

$$0 = \int_{U \times V} \{-f(x) + g(y)\} d\gamma.$$

Since $d\gamma$ has dx, dy as its marginals, it is equivalent to the balance condition

$$(1.10) \qquad \int_U f(x)dx = \int_V g(y)dy.$$

That is to say, in the case of (1.9) the condition (1.8), or equivalently (1.10), is independent of the functions u, v.

As a consequence of Theorem 1.1, we have

Corollary 1.1. *Assume that the function F is given by (1.9) for some positive functions $f \in L^1(U), g \in L^1(V)$ satisfying the balance condition (1.10). Assume that the constraint φ satisfies (1.5) and (H1)–(H2). Then the same conclusions in Theorem 1.1 hold.*

We point out that for the nonlinear optimization (1.1)–(1.2), there is generally no uniqueness for maximizing pair (u, v), see Remark 2.1. Under the following convexity hypothesis

(Hc) φ is convex in s, and F is strictly concave in (t, s)-variable,

one can easily obtain the uniqueness of maximizing pair. However, this hypothesis (Hc) never holds in the case of (1.9). The proof of uniqueness under (Hc) is straight-forward. Suppose $(\bar{u}, \bar{v}) \in K$ and $(\tilde{u}, \tilde{v}) \in K$ are two different maximizing pair, namely

$$(1.11) \qquad I(\bar{u}, \bar{v}) = I(\tilde{u}, \tilde{v}) = \sup_K I =: I^*.$$

Set $\hat{u} = \frac{1}{2}(\bar{u} + \tilde{u})$ and $\hat{v} = \frac{1}{2}(\bar{v} + \tilde{v})$. Since φ is convex in s, we have

$$(1.12) \qquad \phi(x, y, \hat{u}, \hat{v}) \leq \frac{1}{2} \left(\phi(x, y, \bar{u}, \bar{v}) + \phi(x, y, \tilde{u}, \tilde{v}) \right) \leq 0, \ \forall (x, y) \in U \times V,$$

which implies that $(\hat{u}, \hat{v}) \in K$. On the other hand, since F is strictly concave in (t, s),

$$(1.13) \qquad I(\hat{u}, \hat{v}) > \frac{1}{2} \left(I(\bar{u}, \bar{v}) + I(\tilde{u}, \tilde{v}) \right) = I^*,$$

which is a contradiction to (1.11).

It is well known that many constrained nonlinear optimization problems can be solved by Lagrangian dual methods (see, for example, [10]), where the convexity plays a crucial role. For the nonlinear optimization (1.1)–(1.2) under conditions (1.18)–(1.19), which is weaker than the hypothesis (Hc), we can show that there is no duality gap and there exists at least one Lagrange multiplier. This enables us to use the Lagrangian duality theory to study the maximization of the functional I, see Section 4.

Now, let us introduce some terminology in Lagrangian duality. More details are contained in Section 4. Denote $X := C(U) \times C(V)$. The nonlinear optimization (1.1)–(1.2) is equivalent to the following *primal problem*:

$$\text{maximize } I(u, v),$$

$$(1.14) \qquad \text{subject to } (u, v) \in X, \quad \psi(u, v) := \inf_{U \times V} -\phi(x, y, u, v) \geq 0.$$

We always assume that ϕ has the form (1.4). An element $(u, v) \in X$ is called *feasible* if $\psi(u, v) \geq 0$, or equivalently, $(u, v) \in K$.

Define the *Lagrangian function* $L : X \times \mathbb{R} \to \mathbb{R}$ to be

$$(1.15) \qquad L(u, v, \mu) = I(u, v) + \mu \psi(u, v),$$

where $\mu \in \mathbb{R}$. The *dual functional* J is defined by

$$(1.16) \qquad J(\mu) = \sup_{(u,v) \in X} L(u, v, \mu),$$

and the *dual problem* is given by

$$\text{minimize } J(\mu)$$
$$(1.17)$$
$$\text{subject to } \mu \geq 0.$$

Regardless of the functional I and the constraint ϕ of the primal problem, the dual problem has a very nice convexity property, as shown in Lemma 4.2. In the language of nonlinear programming [3, 10], when $\inf_{\mu \geq 0} J(\mu) = \sup_{(u,v) \in K} I(u, v)$, we say that *there is no duality gap*, otherwise, *there is duality gap*.

Theorem 1.2. *Assume that the function F in (1.1) is concave in (t, s), namely for any $(x, y, t, s) \in U \times V \times \mathbb{R}^2$,*

$$(1.18) \qquad Hess_{t,s} F := \begin{pmatrix} \partial_{tt} F, & \partial_{ts} F \\ \partial_{st} F, & \partial_{ss} F \end{pmatrix} \leq 0,$$

and the constraint φ in (1.4) is convex in s, namely for any $(x, y, s) \in U \times V \times \mathbb{R}$,

$$(1.19) \qquad \partial_{ss} \varphi \geq 0.$$

Suppose that there exists a pair $(\bar{u}, \bar{v}) \in X$ such that

$$(1.20) \qquad\qquad \psi(\bar{u}, \bar{v}) > 0.$$

Then there is no duality gap between the primal problem (1.14) and dual problem (1.17), and there exists at least one Lagrange multiplier, (see Definition 4.1).

Note that in the special case (1.9), the convexity assumptions (1.18) and (1.19) imply that $\varphi = \varphi(x, y, v)$ is linear in v, namely

$$\varphi(x, y, v) = c_0(x, y) + c_1(x, y)v,$$

for some functions c_0, c_1, where $c_1 \geq \theta_0 > 0$ in $U \times V$ by the monotonicity (1.5). When $c_1 \equiv 1$, it is an optimal transportation problem with the associated cost function $-c_0$, see Example 5.1.

This paper is organized as follows: In Section 2 we show various properties of potential functions and optimal mappings under appropriate assumptions. In particular, we show the optimal mapping T is measure preserving in the sense of (2.7) and the potential function satisfies a Monge-Ampère type equation. Theorem 1.1 is proved. In Section 3 we introduce a notion of generalized solutions and show that a potential function is a generalized solution, from which the existence of generalized solutions follows, see Theorem 3.1. Section 4 contains the Lagrangian duality theory. It provides a useful tool to obtain the existence of maximizer. Theorem 1.2 is proved. In Section 5 we present some examples and applications of nonlinear optimization (1.1)–(1.2). In particular, we derive the equations in geometric optics problems from the corresponding constraints, instead of using the reflection or refraction law.

2. Existence of maximizers

In this section, we study the existence of potential functions and optimal mappings in the nonlinear optimization problem (1.1)–(1.2).

Lemma 2.1. *Assume that the function F and the constraint ϕ satisfy the hypotheses (1.3)–(1.5) and (1.8). Then the objective functional I in (1.1) has a maximizing pair $(\bar{u}, \bar{v}) \in K$, where K is the constraint set (1.2).*

Proof. Given any pair $(u, v) \in K$, we claim that $I(u, v)$ does not decrease if v is replaced by

(2.1) $$v^*(y) = \sup\{s : \phi(x, y, u(x), s) \leq 0, \quad \forall x \in U\}.$$

In fact, by the continuity of ϕ and u, for each $y \in V$ there is some $x \in \overline{U}$ such that

$$\phi(x, y, u(x), v^*(y)) = 0 \geq \phi(x, y, u(x), v(y)),$$

since $(u, v) \in K$. By (1.4)–(1.5), $v^* \geq v$. Furthermore, $\phi(x, y, u(x), v^*(y)) \leq 0$ for all $(x, y) \in U \times V$, so $(u, v^*) \in K$.

Since $v^* \geq v$, by (1.3) we have

$$I(u, v^*) \geq I(u, v).$$

Similarly, if we define

(2.2) $$u^*(x) = \sup\{t : \phi(x, y, t, v^*(y)) \leq 0, \quad \forall y \in V\},$$

then $(u^*, v^*) \in K$ and

$$I(u^*, v^*) \geq I(u, v^*) \geq I(u, v).$$

Thus we do not decrease $I(u, v)$ by replacing (u, v) by (u^*, v^*). The claim is proved.

Define $K_{C_0} = K \cap \{u \geq C_0\}$, where C_0 is a constant, which may be chosen negative and sufficiently small in the following context. We show that u^* and v^* are uniformly bounded if $(u, v) \in K_{C_0}$. Since $v^* \geq v, u \geq C_0$, by (1.4)–(1.5) we have for each $y \in V$, $s := v^*(y)$,

$$C_0 + \varphi(x, y, s) \leq u(x) + \varphi(x, y, s) \leq 0, \quad \text{for all } x \in U.$$

Then by (1.5) again, there exists a constant C_1, such that $s \leq C_1$. This implies that

(2.3) $$v \leq v^* \leq C_1,$$

we may choose C_1 such that $\sup_V v^* = C_1$. By a similar argument, there is another constant $\tilde{C}_0$ depending on φ and C_1 such that $\inf_U u^* = \tilde{C}_0$. The constant $\tilde{C}_0 \geq C_0$, since $u^* \geq u$ in U, and so $(u^*, v^*) \in K_{C_0}$.

We next deduce the lower bound of v^* and the upper bound of u^* by showing that u^* and v^* are locally Lipschitz. Consider two points in U, $x_1 \neq x_2$ and $|x_1 - x_2| < \varepsilon$ sufficiently small. There are two points $y_1, y_2 \in \overline{V}$ such that

$$\phi(x_1, y_1, u^*(x_1), v^*(y_1)) = 0,$$
$$\phi(x_2, y_2, u^*(x_2), v^*(y_2)) = 0.$$

Then by (1.4), we have

$$\begin{aligned}
0 =& \phi(x_2, y_2, u^*(x_2), v^*(y_2)) - \phi(x_1, y_2, u^*(x_1), v^*(y_2)) \\
& + \phi(x_1, y_2, u^*(x_1), v^*(y_2)) - \phi(x_1, y_1, u^*(x_1), v^*(y_1)) \\
=& u^*(x_2) - u^*(x_1) - \varphi_x(\hat{x}, y_2, v^*(y_2)) \cdot (x_2 - x_1) \\
& + \phi(x_1, y_2, u^*(x_1), v^*(y_2)),
\end{aligned}$$

where $\hat{x} = \theta x_1 + (1 - \theta)x_2$ for some $\theta \in (0, 1)$. Noting that $\phi(x_1, y_2, u^*(x_1), v^*(y_2)) \leq 0$, we have

$$u^*(x_2) - u^*(x_1) \geq -C_2|x_2 - x_1|,$$

where the constant $C_2 = \sup(|\varphi_x| + |\varphi_y|)$.

On the other hand, replacing $\phi(x_1, y_2, u^*(x_1), v^*(y_2))$ by $\phi(x_2, y_1, u^*(x_2), v^*(y_1))$ in the above calculation, we have

$$u^*(x_2) - u^*(x_1) \leq C_2|x_2 - x_1|.$$

Therefore, the Lipschitz constant of u^* on U is controlled by

$$(2.4) \qquad \qquad \|u^*\|_{Lip(U)} \leq C_2.$$

By switching x and y in the above argument, we can obtain the Lipschitz continuity of v^* on V,

$$\begin{aligned}
|v^*(y_2) - v^*(y_1)| &\leq \frac{\sup|\varphi_y|}{\inf \varphi_s}|y_2 - y_1| \\
&\leq C_2\theta_0^{-1}|y_2 - y_1|,
\end{aligned}$$

where θ_0 is the constant in (1.5), y_1, y_2 are two distinct points in V. This inequality implies that $\|v^*\|_{Lip(V)} \leq C_2\theta_0^{-1}$. Hence, we have $u^* \leq \tilde{C}_0 + C_2\text{diam}(U)$ and $v^* \geq C_1 - C_2\theta_0^{-1}\text{diam}(V)$.

We conclude, therefore, that any pair $(u, v) \in K_{C_0}$ may be replaced by a bounded, Lipschitz pair $(u^*, v^*) \in K_{C_0}$ without decreasing I. We now choose a maximizing sequence $\{(u_k, v_k)\} \subset K_{C_0}$ such that

$$I(u_k, v_k) \to \sup_{(u,v) \in K_{C_0}} I(u, v).$$

By the above considerations we may assume that each (u_k, v_k) is a bounded, uniformly Lipschitz pair, uniformly with respect to k, so there is a subsequence converging uniformly to a bounded, Lipschitz, maximizing pair $(\bar{u}, \bar{v}) \in K_{C_0}$.

Last, we show that when $C_0 < 0$ sufficiently small,

$$\sup_{(u,v) \in K_{C_0}} I(u, v) = \sup_{(u,v) \in K} I(u, v),$$

or equivalently, $\sup_{K_{C_0}} I$ is independent of C_0. By definition, one has $\sup_{K_{C_0-1}} I \geq \sup_{K_{C_0}} I$. So it suffices to show the reverse inequality. Let $(u, v) \in K_{C_0-1}$ be a maximizer such that $I(u, v) = \sup_{K_{C_0-1}} I$, and $\{x_k\}_{k=1,\cdots,N}$ be a set of points in U. For a small constant $\varepsilon > 0$, define

$$\tilde{u} = \begin{cases} u & \text{in } U - \cup_N B_\varepsilon(x_k), \\ u+2 & \text{in } \cup_N B_\varepsilon(x_k). \end{cases}$$

Note that we may replace $\tilde{u}$ by its mollification $\tilde{u}_h = \rho_h * \tilde{u}$, where ρ_h is the standard mollifier function [9]. For simplicity, we assume $\tilde{u}$ continuous in the sense that for $h > 0$ sufficiently small,

$$I(\tilde{u}_h, v) = I(u, v) + O(N\varepsilon^n).$$

Define

$$\tilde{v}^*(y) = \sup\{s : \phi(x, y, \tilde{u}(x), s) \leq 0, \quad \forall x \in U\},$$
$$\tilde{u}^*(x) = \sup\{t : \phi(x, y, t, \tilde{v}^*(y)) \leq 0, \quad \forall y \in V\}.$$

Since the constraint function φ is C^1 smooth in s and by (1.4)–(1.5), except a set $E \subset U$ and a set $E' \subset V$ of measure $|E| = |E'| = O(N\varepsilon^n)$,

$$\tilde{v}^* = v - \frac{2}{\varphi_s} + O(\delta) \quad \text{in } V \setminus E',$$
$$\tilde{u}^* = u + 2 + O(\delta) \quad \text{in } U \setminus E,$$

where $\delta := \min_{i \neq j}\{\text{dist}(x_i, x_j)\}$. Therefore, by (1.8) and the mean value theorem we have

$$I(\tilde{u}^*, \tilde{v}^*) = I(u, v) + 2 \int_{(U \setminus E) \times (V \setminus E')} \left\{ F_t - \frac{F_s}{\varphi_s} \right\} d\gamma + O(\delta) + O(N\varepsilon^n)$$
$$\geq I(u, v) - C\delta - CN\varepsilon^n.$$

As $(u, v) \in K_{C_0-1}$, we may assume that $\inf_U u = C_0 - 1$. Otherwise, one has $\inf_U u = C_0 - \tau_0$ for some constant $\tau_0 < 1$. This implies that $\sup_{K_{C_0-1}} I = \sup_{K_{C_0}} I$, namely $\sup_{K_{C_0}} I$ is independent of C_0, and the proof is finished. By the definition, δ will become small if the number of points N is sufficiently large so that we have $(\tilde{u}^*, \tilde{v}^*) \in K_{C_0}$ and

$$\sup_{K_{C_0}} I \geq I(\tilde{u}^*, \tilde{v}^*) \geq \sup_{K_{C_0-1}} I - C\delta - CN\varepsilon^n.$$

Then, choosing $\varepsilon > 0$ sufficiently small, we have

$$\sup_{K_{C_0-1}} I \leq \sup_{K_{C_0}} I,$$

by letting $\delta \to 0, \varepsilon \to 0$, which implies that $\sup_{K_{C_0}} I$ is independent of C_0, and the proof is finished. □

Remark 2.1. From the proof of Lemma 2.1, we conclude that there exist infinitely many maximizing pairs. In fact, if (u, v) is a maximizer and $C_0 = \inf_U u$, then there is another maximizer in K_{C_0+1}, which is different from (u, v).

Lemma 2.2. *Let $(u, v) \in K$ be a dual maximizing pair of I in Lemma 2.1. Assume φ satisfies condition (H1). There exists an associated mapping T with respect to (u, v), which solves equation* (1.7) *and is uniquely determined almost everywhere on U.*

Proof. Since u satisfies (1.6) and v, φ are continuous, for each $x \in U$, there exists some $y =: T(x) \in \overline{V}$ such that

$$(2.5) \qquad\qquad u(x) + \varphi(x, y, v(y)) \;=\; 0,$$
$$u(x') + \varphi(x', y, v(y)) \;\leq\; 0,$$

for any other $x' \in U$. Note that u is differentiable almost everywhere since u is Lipschitz continuous. Let $x \in U$ be a differentiable point of u, by differentiation we have

$$(2.6) \qquad\qquad \varphi_x(x, y, v) + Du(x) = 0.$$

Therefore, for the fixed $x \in U$, setting $t = u(x)$ and $p = Du(x)$ one can see that

$$\varphi_x(x, y, v) = -p, \quad \text{and } \varphi(x, y, v) = -t.$$

From the condition (H1), we then obtain the uniqueness of $y = T(x)$. Since u is differentiable almost everywhere on U, the mapping T is determined almost everywhere on U. □

Remark 2.2. In the condition (H1), we assume that $(\varphi_x, \varphi)(x, \cdot, \cdot)$ is one-to-one in the whole space $\mathbb{R}^n \times \mathbb{R}$, for each $x \in U$. This is only for simplicity. We may allow that the constraint φ is defined in a proper subset $\mathcal{U} \times \mathcal{I}$, where $\mathcal{U} \subset \mathbb{R}^n \times \mathbb{R}^n$ and $\mathcal{I} \subset \mathbb{R}$. Denote the projections $\mathcal{U}_x = \{y \in \mathbb{R}^n : (x, y) \in \mathcal{U}\}$, $\mathcal{U}_y = \{x \in \mathbb{R}^n : (x, y) \in \mathcal{U}\}$, let $U = \bigcup \mathcal{U}_y$, $V = \bigcup \mathcal{U}_x$. In this case we replace (H1) by assuming that: for any $(x, y) \in \mathcal{U}$, there exists an open interval $\mathcal{I}(x, y) \subset \mathcal{I}$, such that $(\varphi_x, \varphi)(x, \cdot, \cdot)$ is one-to-one in $y \in \mathcal{U}_x, s \in \mathcal{I}(x, y)$, for each $x \in U$. Accordingly the condition (H2) can be restricted to $\mathcal{U} \times \mathcal{I}$, namely for any $(x, y) \in \mathcal{U}$, $s \in \mathcal{I}(x, y)$, we have $\det(\varphi_{xy} - \varphi_s^{-1}\varphi_{xs} \otimes \varphi_y) \neq 0$.

Let $(u, v) \in K$ be a dual maximizing pair of I, and $S : U \to V$. We say the mapping S is *measure preserving* if

$$(2.7) \qquad 0 = \int_{U \times V} \{-F_t \varphi_s(x, Sx, v(Sx))h(Sx) + F_s(x, y, u(x), v(y))h(y)\}\, d\gamma,$$

for any $h \in C(V)$.

Lemma 2.3. *Assume that the balance condition* (1.8) *holds. Let T be the optimal mapping obtained in Lemma 2.2, associated with a dual maximizing pair (u, v). Then T is measure preserving in the sense of* (2.7).

Proof. Let $h \in C(V)$ and $|\epsilon| < 1$ sufficiently small. Define

$$(2.8) \qquad\qquad v_\epsilon(y) = v(y) + \epsilon h(y)$$

and

$$(2.9) \qquad u_\epsilon(x) = \sup\{t \; : \; t + \varphi(x, y, v_\epsilon(y)) \le 0, \quad \forall y \in V\}.$$

Then $(u_\epsilon, v_\epsilon) \in K$ and $(u_0, v_0) = (u, v)$.

Since (u, v) satisfies (1.6), by Lemma 2.2 for every $x \in U$ the supremum (1.6) is attained at point $y_0 = T(x)$. We claim that at these points we have

$$(2.10) \qquad u_\epsilon(x) - u(x) = -\epsilon\varphi_s h(T(x)) + o(\epsilon).$$

To prove (2.10), first we show that $LHS \le RHS$.

$$
\begin{aligned}
0 &= u(x) + \varphi(x, y_0, v(y_0)) \\
&= u(x) + \varphi(x, y_0, v_\epsilon(y_0) - \epsilon h(y_0)) \\
&= u(x) + \varphi(x, y_0, v_\epsilon(y_0)) - \epsilon\phi_s h(y_0) + o(\epsilon) \\
&= u(x) - u_\epsilon(x) + \phi(x, y_0, u_\epsilon(x), v_\epsilon(y_0)) - \epsilon\varphi_s h(y_0) + o(\epsilon) \\
&\le u(x) - u_\epsilon(x) - \epsilon\varphi_s h(y_0) + o(\epsilon).
\end{aligned}
$$

Thus we have

$$u_\epsilon(x) - u(x) \le -\epsilon\varphi_s h(y_0) + o(\epsilon).$$

To show $LHS \ge RHS$ we use the fact that for any such $x \in U$ there are points $y_\epsilon \in \overline{V}$ such that the supremum in (2.9) is attained. Thus

$$
\begin{aligned}
0 &\ge u(x) + \varphi(x, y_\epsilon, v(y_\epsilon)) \\
&= u(x) + \varphi(x, y_\epsilon, v_\epsilon(y_\epsilon) - \epsilon h(y_\epsilon)) \\
&= u(x) + \varphi(x, y_\epsilon, v_\epsilon(y_\epsilon)) - \epsilon\varphi_s h(y_\epsilon) + o(\epsilon) \\
&= u(x) - u_\epsilon(x) + \phi(x, y_\epsilon, u_\epsilon(x), v_\epsilon(y_\epsilon)) - \epsilon\varphi_s h(y_\epsilon) + o(\epsilon) \\
&= u(x) - u_\epsilon(x) - \epsilon\varphi_s h(y_\epsilon) + o(\epsilon).
\end{aligned}
$$

Then we have

$$
\begin{aligned}
u_\epsilon(x) - u(x) &\ge -\epsilon\varphi_s h(y_\epsilon) + o(\epsilon) \\
&= -\epsilon\varphi_s h(y_0) + \epsilon\varphi_s \left(h(y_0) - h(y_\epsilon)\right) + o(\epsilon).
\end{aligned}
$$

Since the supremum in (1.6) is attained at y_0, we have $y_\epsilon \to y_0$ as $\epsilon \to 0$, and therefore, since $h \in C(V)$,

$$\epsilon\varphi_s \left(h(y_0) - h(y_\epsilon)\right) = o(\epsilon).$$

This implies that $LHS \ge RHS$, and (2.10) follows.

Next, since $(u, v) = (u_0, v_0)$ maximizes I among all pairs in K, we obtain

$$
\begin{aligned}
0 &= \lim_{\epsilon \to 0} \frac{I(u_\epsilon, v_\epsilon) - I(u, v)}{\epsilon} \\
&= \lim_{\epsilon \to 0} \int_{U \times V} \frac{F(x, y, u_\epsilon, v_\epsilon) - F(x, y, u, v)}{\epsilon} d\gamma \\
&= \int_{U \times V} \{-F_t \varphi_s h(T(x)) + F_s h(y)\} \, d\gamma.
\end{aligned}
$$

□

In the special case of (1.9), by (1.5) and the arbitrariness of test function $h \in C(V)$ in (2.7) we have

$$
(2.11) \qquad \int_U f(x) h(Sx) dx = \int_V g(y) h(y) dy,
$$

for any $h \in C(V)$, where f, g are positive functions in $L^1(U), L^1(V)$, respectively, and satisfy the balance condition (1.10). Moreover, we have the following

Corollary 2.1. *Assume the function F is given by (1.9) with $f > 0, \in L^1(U)$, $g > 0, \in L^1(V)$ satisfying (1.10). The optimal mapping T in Lemma 2.2 is measure preserving in the sense of (2.11). If T is continuous differentiable, then*

$$
(2.12) \qquad |det\, DT| = \frac{f}{g \circ T}.
$$

Proof. From the proof of Lemma 2.3, it is obvious that T is measure preserving in the sense of (2.11). When T is C^1 smooth, by the formula of change of coordinates,

$$
\int_U f(x) h(Tx) dx = \int_U g(Tx) h(Tx) |det\, DT| dx,
$$

for any $h \in C(V)$. Hence the Jacobian of DT satisfies (2.12). □

As a consequence of Lemma 2.3 and Corollary 2.1, we derive the equation satisfied by the potential function u as follows. At this stage, let us assume all the functions are smooth enough, say at least C^2, so that we can do the differentiations.

Let $(u, v) \in K$ be the dual maximizing pair of I, and T be the associated optimal mapping. By Definition 1.2 and (2.6) we have

$$
\varphi_x(x, Tx, v(Tx)) + Du(x) = 0,
$$

in U. Differentiating with respect to x, we then get

$$
(2.13) \qquad 0 = \varphi_{xx} + \varphi_{xy} DT + (\varphi_{xs} \otimes Dv) DT + D^2 u,
$$

where each side is regarded as an $n \times n$ matrix valued at (x, y), $y = T(x)$.

In order to eliminate Dv in (2.13), we note that for $x_0 \in U$ equality (1.7) holds at $y_0 = T(x_0)$, and for other $y' \in V$

$$u(x_0) + \varphi(x_0, y', v(y')) \leq 0,$$

since $(u, v) \in K$. Thus, at (x_0, y_0) there holds

$$\frac{d\varphi}{dy} = \varphi_y + \varphi_s Dv = 0.$$

By the assumption (1.5), $\varphi_s > 0$ we thus get

(2.14) $$Dv = -\frac{\varphi_y}{\varphi_s}.$$

Combining (2.13) and (2.14) we have the equation

(2.15) $$\left| D^2 u + \varphi_{xx} \right| = \left| \varphi_{xy} - \frac{1}{\varphi_s} \varphi_{xs} \otimes \varphi_y \right| |DT|.$$

In the case of (1.9)–(1.10), by Corollary 2.1 we obtain the equation

(2.16) $$\left| \det \left[D^2 u + \varphi_{xx} \right] \right| = \left| \det \left[\varphi_{xy} - \frac{1}{\varphi_s} \varphi_{xs} \otimes \varphi_y \right] \right| \frac{f}{g \circ T}.$$

Under the condition (H2) and $f, g > 0$ one see that (2.16) is a non-degenerate Monge-Ampère type equation [9, 11]. When the optimal mapping T is a diffeomorphism from U to V, we have the natural boundary condition

(2.17) $$T(U) = V.$$

Similarly, one can also derive the dual PDE for the dual potential v.

When $\varphi(x, y, v) = x \cdot y$, then (2.16) is equivalent to the standard Monge-Ampère equation

$$\det D^2 u = h,$$

with the boundary condition

$$Du(U) = V,$$

where $h = f/g$. When $\varphi(x, y, v) = v - c(x, y)$ for some function $c : U \times V \to \mathbb{R}$, we have the optimal transportation equation, see Example 5.1,

$$\det \left[D^2 u - D_x^2 c \right] = |\det D_{xy}^2 c| \frac{f}{g \circ T},$$

with the boundary condition

$$T(U) = V.$$

It is well-known that the regularity of equation (2.16) depends crucially on the structure of constraint function φ, as in [19, 24]. In the next section, we introduce a notion of a generalized solution and show that if (u, v) is a dual maximizing pair of I over K, then the potential u is a generalized solution of (2.16). The regularity property of u will be investigated in a separate paper [17].

3. Generalized solution

In this section, we introduce a notion of generalized solutions of (2.16) and show that the potential function is a generalized solution. Let φ be the constraint in (1.4). First we introduce the φ-concavity for functions, which is an extension of the c-concavity in optimal transportation, see [8, 19].

Definition 3.1. *A φ-support function of u at x_0 is a function of the form $\varphi(x, y_0, s_0)$, where $y_0 \in \mathbb{R}^n$, and $s_0 \in \mathbb{R}$ is a constant such that*

$$u(x_0) + \varphi(x_0, y_0, s_0) = 0,$$

(3.1)
$$u(x) + \varphi(x, y_0, s_0) \leq 0 \quad \forall x \in U.$$

A continuous function u defined on $\overline{U}$ is φ-concave if for any point $x_0 \in U$, there exists a φ-support function at x_0.

By definition, the potential function u is φ-concave with $y_0 \in V$, $s_0 = v(y_0)$. In the special case when $\varphi(x, y, s) = s - x \cdot y$, the notion of φ-concavity coincides with that of concavity, and the graph of a φ-support function is a support hyperplane.

Recall that φ is derived from the constraint function $\phi(x, y, u, v)$ by the strict monotonicity in u. Since ϕ is also strictly increasing in v, the constraint (1.4) can also be written as

(3.2)
$$\phi(x, y, u, v) = v + \varphi^*(x, y, u) \leq 0,$$

for a function $\varphi^* = \varphi^*(x, y, t)$ strictly increasing in t. The function $\varphi^* = \varphi^*(x, y, t)$ is called *dual constraint function* of φ in the sense of

$$-\varphi(x, y, -\varphi^*(x, y, t)) = t,$$
$$-\varphi^*(x, y, -\varphi(x, y, s)) = s,$$

for all $(x, y) \in U \times V$. By differentiating, we have

(3.3)
$$\varphi_t^* = \frac{1}{\varphi_s}, \quad \varphi_x^* = \frac{\varphi_x}{\varphi_s}, \quad \varphi_y^* = \frac{\varphi_y}{\varphi_s}.$$

For the dual constraint φ^*, we assume the following condition analogous to (H1):

(H1*) For each $y_0 \in V$, for any $(q, s) \in \mathbb{R}^n \times \mathbb{R}$, there is at most one pair $(x, t) \in \mathbb{R}^n \times \mathbb{R}$ such that

$$\left(\varphi_y^*, \varphi^*\right)(x, y_0, t) = -(q, s),$$

namely, $\varphi_y^*(x, y_0, t) = -q$ and $\varphi^*(x, y_0, t) = -s$.

The φ-concavity in Definition 3.1 and (3.1)–(3.3) are generalizations of c-concavity and c-duality in optimal transportation, where

$$\varphi(x, y, s) = s - c(x, y), \quad \varphi^*(x, y, t) = t - c(x, y),$$

for a cost function $c(x, y)$. The condition (H1)–(H1*) is the counterpart of the condition (A1) assumed on the cost function $c(x, y)$ in [19]. Note that from (3.2) and (3.3), we can directly derive (2.14) for a dual pair of functions u, v.

Similarly, by switching x and y, U and V, one can also introduce the notion of φ^*-concavity for the function v. From Definition 1.1 and (3.1), when $(u, v) \in K$ is a dual pair, u is naturally φ-concave and v is φ^*-concave.

Let u be a φ-concave function in U. We define a set-valued mapping $T_u = T_{u,\varphi} : U \to V$. For any $x_0 \in U$, let $T_u(x_0)$ denote the set of points y_0 such that $\varphi(x, y_0, s_0)$ is a φ-support function of u at x_0 for some constant s_0. For any subset $E \subset U$, we denote $T_u(E) = \cup_{x \in E} T_u(x)$.

If u is C^1 smooth, by condition (H1) (y_0, s_0) is uniquely determined by $(Du(x_0), u(x_0))$, and T_u is single valued. In this paper we call the mapping T_u the φ-*normal mapping* of u. Similarly we can define the φ^*-normal mapping for φ^*-concave functions. In particular, if $(u, v) \in K$ is a dual pair, we see that $y \in T_{u,\varphi}(x)$ if and only if $x \in T_{v,\varphi^*}(y)$.

Remark 3.1. As the constraint function φ is smooth, any φ-concave function u is semi-concave, namely there exists a constant C such that $u(x) - C|x|^2$ is concave. It follows that u is twice differentiable almost everywhere and $T_u(x)$ is a singleton for almost all $x \in U$.

Lemma 3.1. *Let $(u, v) \in K$ be a dual maximizing pair of I. Assume that the constraint φ^* satisfies condition (H1*). Let*

$$Y = Y_u = \{y \in V \mid \exists\, x_1 \neq x_2 \in U \text{ such that } y \in T_u(x_1) \cap T_u(x_2)\}.$$

Then Y has Lebesgue measure zero.

Proof. If $y \in T_u(x_1) \cap T_u(x_2)$, we have $x_1, x_2 \in T_{v,\varphi^*}(y)$. From the proof of Lemma 2.1, v is almost everywhere differentiable. Assume y is a differentiable point, then by definition

$$\varphi^*(x_i, y, u(x_i)) = -v(y),$$
$$\varphi_y^*(x_i, y, u(x_i)) = -Dv(y),$$

for $i = 1, 2$. If $x_1 \neq x_2$, this is a contradiction to (H1*). $\qquad\square$

We now define a measure $\mu = \mu_{u,g}$ in U, where $g \in L^1(V)$ is the positive measurable function in (1.9)–(1.10). Set $g \equiv 0$ in $\mathbb{R}^n - V$. For any Borel set $E \subset U$, define

$$(3.4) \qquad\qquad \mu(E) = \int_{T_u(E)} g(y)dy.$$

It follows from Lemma 3.1 that μ is a Radon measure, and satisfies the following regularity properties:

$$\mu(E) = \inf\{\mu(D) : E \subset D \subset U,\, D \text{ open}\}$$

for all Borel sets $E \subset U$, and

$$\mu(D) = \sup\{\mu(K) \ : \ K \subset D, \ K \text{ compact}\}$$

for all open sets $D \subset U$. For further discussion of the measure μ and its stability property, see [2, 6, 11, 19].

Definition 3.2. *A φ-concave function u is called a generalized solution of* (2.16) *if $\mu_{u,g} = f dx$ in the sense of measure, that is for any Borel set $E \subset U$,*

(3.5)
$$\int_E f = \int_{T_u(E)} g.$$

Note that since we extended $g = 0$ to $\mathbb{R} - V$, the boundary condition (2.17) is a consequence of the mass balance condition (1.10) in the sense of $|V - T_u(U)| = 0$.

The next result shows that a potential function u is a generalized solution. The existence of potential in Lemma 2.1 then implies the existence of generalized solutions. But in general we do not have the uniqueness, see Remark 2.1.

Theorem 3.1. *Let $(u,v) \in K$ be a dual maximizing pair of I. Then u is a generalized solution of* (2.16).

Proof. Let $(u,v) \in K$ be a dual maximizing pair of I. Then by (1.6), u is φ-convex and v is φ^*-convex with respect to each other. By Lemma 2.2, the optimal mapping T associated to (u,v), as determined by (2.6), is equal to the mapping $T_{u,\varphi}$ almost everywhere on U. By Corollary 2.1 T is measure preserving in the sense of (2.11). Hence u is a generalized solution of (2.16). Assumption (1.10) implies that (2.17) holds. $\square$

4. Lagrangian duality

In this section, we study the dual problem (1.17) of the constrained nonlinear optimization (1.1)–(1.2), and prove Theorem 1.2. Recall that the Lagrangian function L is defined in (1.15), where the constraint ψ is given in (1.14). Denote by I^* the optimal value of the primal problem (1.14), namely

$$I^* = \sup_{(u,v) \in K} I(u,v).$$

Definition 4.1. *A factor μ^* is called a* Lagrange multiplier *for the primal problem if $\mu^* \geq 0$, and*

$$I^* = \sup\{L(u,v,\mu^*) \ : \ (u,v) \in X\}.$$

Lemma 4.1. *Let μ^* be a Lagrange multiplier. Then (u^*, v^*) is a global maximum of the primal problem if and only if (u^*, v^*) is feasible and*

$$(4.1) \qquad\qquad (u^*, v^*) = \arg \max_{(u,v) \in X} L(u, v, \mu^*),$$

$$(4.2) \qquad\qquad \mu^* \psi(u^*, v^*) = 0.$$

Proof. If (u^*, v^*) is a global maximum of the primal problem, then (u^*, v^*) is feasible and furthermore,

$$
\begin{aligned}
I^* = I(u^*, v^*) &\leq I(u^*, v^*) + \mu^* \psi(u^*, v^*) \\
(4.3) \qquad &= L(u^*, v^*, \mu^*) \leq \sup\{L(u, v, \mu^*) \, : \, (u, v) \in C(U) \times C(V)\} \\
&= I^*,
\end{aligned}
$$

where the first inequality follows from the definition of Lagrange multiplier ($\mu^* \geq 0$) and the feasibility of (u^*, v^*) (i.e. $\psi(u^*, v^*) \geq 0$). Using again the definition of Lagrange multiplier, we have $I^* = \sup_{(u,v) \in X} L(u, v, \mu^*)$, so that equality holds throughout (4.3). This implies the equalities (4.1)–(4.2).

Conversely, if (u^*, v^*) is feasible and (4.1)–(4.2) hold, we have from the definition of Lagrange multiplier,

$$
\begin{aligned}
I(u^*, v^*) &= I(u^*, v^*) + \mu^* \psi(u^*, v^*) \\
&= L(u^*, v^*, \mu^*) = \max_{(u,v) \in X} L(u, v, \mu^*) = I^*,
\end{aligned}
$$

so (u^*, v^*) is a global maximum. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Recall the definitions of the dual functional J in (1.16) and the dual problem in (1.17). Note that $J(\mu)$ may be equal to $+\infty$ for some μ. In this case, we define the domain of J to be the set of μ for which $J(\mu)$ is finite:

$$D = \{\mu \in \mathbb{R} \, : \, J(\mu) < +\infty\}.$$

Regardless of the functional I and the constraint ϕ of the primal problem, the dual problem (1.17) has a nice convexity property, as shown in the following lemma.

Lemma 4.2. *The domain D of the dual functional J is convex and J is convex over D.*

Proof. For any $u, v, \mu, \bar{\mu}$, and $\alpha \in [0, 1]$, we have

$$L(u, v, \alpha\mu + (1 - \alpha)\bar{\mu}) = \alpha L(u, v, \mu) + (1 - \alpha)L(u, v, \bar{\mu}).$$

Taking the supremum over all $(u, v) \in X$, we obtain

$$\sup L(u, v, \alpha\mu + (1 - \alpha)\bar{\mu}) \leq \alpha \sup L(u, v, \mu) + (1 - \alpha) \sup L(u, v, \bar{\mu}),$$

or equivalently

$$J(\alpha\mu + (1 - \alpha)\bar{\mu}) \leq \alpha J(\mu) + (1 - \alpha)J(\bar{\mu}).$$

Therefore if μ and $\bar{\mu}$ belong to D, the same is true for $\alpha\mu + (1 - \alpha)\bar{\mu}$, so D is convex. Furthermore, J is convex over D. □

Another important property is that the optimal dual value

$$J^* = \inf_{\mu \geq 0} J(\mu)$$

is always an upper bound of the optimal primal value, as shown in the next lemma.

Lemma 4.3. *We have*

$$I^* \leq J^*.$$

Proof. For all $\mu \geq 0$, and $(u, v) \in X$ with $\psi(u, v) \geq 0$, we have

$$J(\mu) = \sup_{(\tilde{u},\tilde{v}) \in X} L(\tilde{u}, \tilde{v}, \mu)$$
$$\geq I(u, v) + \mu\psi(u, v) \geq I(u, v),$$

and therefore,

$$J^* = \inf_{\mu \geq 0} J(\mu) \geq \sup_{(u,v) \in K} I(u, v) = I^*.$$

□

In the language of nonlinear programming [3, 10], if $J^* = I^*$ we say that *there is no duality gap*; if $J^* > I^*$ *there is duality gap*. Note that if there exists a Lagrange multiplier μ^*, the above lemma ($J^* \geq I^*$) and the definition of Lagrange multiplier ($I^* = J(\mu^*) \geq J^*$) imply that there is no duality gap.

The following is a sufficient condition for the existence of Lagrange multiplier, which is also a proof of Theorem 1.2.

Lemma 4.4. *Let the assumptions* (1.18)–(1.20) *hold for the primal problem* (1.14). *Then there is no duality gap and there exists at least one Lagrange multiplier.*

Proof. Consider the subset of $\mathbb{R}^2$ given by

$$A = \{(z, w) : \exists (u, v) \in X \text{ such that}$$
$$\psi(u, v) \geq z, \quad I(u, v) \geq w\}.$$

We first show that A is convex. Let $(z, w) \in A$ and $(\tilde{z}, \tilde{w}) \in A$ be two different elements, we show that their convex combinations belong to A.

The definition of A implies that for some $(u, v) \in X$ and $(\tilde{u}, \tilde{v}) \in X$, we have

$$I(u, v) \geq w, \quad \psi(u, v) \geq z,$$
$$I(\tilde{u}, \tilde{v}) \geq \tilde{w}, \quad \psi(\tilde{u}, \tilde{v}) \geq \tilde{z}.$$

For any $\alpha \in [0,1]$, by the concavity of F in (1.18), we obtain

$$I\left(\alpha u + (1-\alpha)\tilde{u}, \alpha v + (1-\alpha)\tilde{v}\right) \geq \alpha I(u,v) + (1-\alpha)I(\tilde{u},\tilde{v})$$

$$\geq \alpha w + (1-\alpha)\tilde{w}.$$

By the convexity of φ in (1.19) and noting that $\inf_\Omega (f + h) \geq \inf_\Omega f + \inf_\Omega h$ for any $f, h \in C^0(\Omega)$, we obtain

$$\psi\left(\alpha u + (1-\alpha)\tilde{u}, \alpha v + (1-\alpha)\tilde{v}\right) \geq \alpha \psi(u,v) + (1-\alpha)\psi(\tilde{u},\tilde{v})$$

$$\geq \alpha z + (1-\alpha)\tilde{z}.$$

Since the combination $(\alpha u + (1-\alpha)\tilde{u}, \alpha v + (1-\alpha)\tilde{v}) \in X$, the above equations imply that the convex combination of (z,w) and $(\tilde{z},\tilde{w})$, i.e.

$$\left(\alpha z + (1-\alpha)\tilde{z}, \alpha w + (1-\alpha)\tilde{w}\right),$$

belongs to A. This proves the convexity of A.

We next observe that $(0, I^*)$ is not an interior point of A; otherwise, the point $(0, I^* + \varepsilon)$ would belong to A for some small $\varepsilon > 0$, contradicting the definition of I^* as the optimal primal value.

Therefore, there exists a supporting hyperplane passing through $(0, I^*)$ and containing A in one side. In particular, there exists a vector $(\mu, \beta) \neq (0,0)$ such that

$$(4.4) \qquad\qquad \beta I^* \geq \mu z + \beta w, \qquad \forall (z,w) \in A.$$

We observe that if $(z,w) \in A$, then $(z, w - \gamma) \in A$ and $(z - \gamma, w) \in A$ for all $\gamma > 0$. The inequality (4.4) thus implies that

$$(4.5) \qquad\qquad \mu \geq 0, \qquad \beta \geq 0.$$

We now claim that $\beta > 0$. If not, $\beta = 0$ and from (4.4)

$$0 \geq \mu z, \qquad \forall (z,w) \in A.$$

By the assumption (1.20), there exists a pair $(\bar{u}, \bar{v}) \in X$ such that

$$\psi(\bar{u}, \bar{v}) > 0.$$

Since $(\psi(\bar{u}, \bar{v}), I(\bar{u}, \bar{v})) \in A$, we have

$$0 \geq \mu \psi(\bar{u}, \bar{v}),$$

which in view of $\mu \geq 0$ in (4.5) implies that $\mu = 0$. This means, however, that $(\mu, \beta) = (0,0)$ arriving at a contradiction. Thus, we must have $\beta > 0$ and by dividing if necessary the vector (μ, β) by β, we may assume that $\beta = 1$. Note that

$$(\psi(u,v), I(u,v)) \in A, \qquad \forall (u,v) \in X.$$

Equation (4.4) implies that

$$(4.6) \qquad\qquad I^* \geq I(u,v) + \mu \psi(u,v), \qquad \forall (u,v) \in X.$$

Taking the supremum over $(u,v) \in X$ and using the fact $\mu \geq 0$, we obtain

$$I^* \geq \sup_{(u,v) \in X} L(u,v,\mu)$$
$$= J(\mu) \geq J^*,$$

where J^* is the optimal dual value. By Lemma 4.3 we have the equalities hold above, namely μ is a Lagrange multiplier and there is no duality gap. $\square$

5. EXAMPLES AND APPLICATIONS

In this section we present some interesting examples and applications of the nonlinear optimization (1.1)–(1.2).

5.1. **Optimal transportation.** Let U, V be two bounded domains in $\mathbb{R}^n$, and $c \in C^4(U \times V)$ be a cost function. Let f, g be two positive densities supported on U, V, respectively, satisfying the mass balance condition

$$(5.1) \qquad \int_U f = \int_V g.$$

The optimal transportation problem is to find a measure preserving mapping $T_0 : U \to V$ minimizing the cost functional

$$\mathcal{C}(T) = \int_U c(x, T(x)) f(x) dx$$

among all measure preserving mappings T from U to V. A mapping $T : U \to V$ is called measure preserving if

$$\int_U h(T(x)) f(x) dx = \int_V h(y) g(y) dy$$

for any $h \in C(V)$. Denote the set of measure preserving mappings by $\mathcal{T}$.

Kantorovich introduced a dual functional

$$(5.2) \qquad I(u,v) = \int_U u(x) f(x) dx + \int_V v(y) g(y) dy$$

over the set

$$(5.3) \qquad K = \{(u,v) \in C(U) \times C(V) : u(x) + v(y) \leq c(x,y), \quad \forall x \in U, y \in V\}.$$

Under suitable conditions, one can prove that

$$\inf_{T \in \mathcal{T}} \mathcal{C}(T) = \sup_{(u,v) \in K} I(u,v).$$

The reader is referred to [1, 5, 7, 8, 19, 22, 23] for further discussion on the optimal transportation problem.

Note that (5.2)–(5.3) is a linear case of (1.1)–(1.2):

$$F(x,y,u,v) = u(x) f(x) + v(y) g(y), \quad d\gamma = dx \otimes dy,$$
$$\phi(x,y,u,v) = u(x) + v(y) - c(x,y),$$

or equivalently, $\varphi(x, y, v) = v - c(x, y)$ in (1.4) and $\varphi^*(x, y, u) = u - c(x, y)$ in (3.2). All the hypotheses in Theorem 1.1 are satisfied when the cost function $c(x, y)$ satisfies the following conditions [19]:

(A1) For any $x, p \in \mathbb{R}^n$, there is a unique $y \in \mathbb{R}^n$ such that $D_x c(x, y) = p$; and for any $y, q \in \mathbb{R}^n$, there is a unique $x \in \mathbb{R}^n$ such that $D_y c(x, y) = q$.

(A2) For any $x, y \in \mathbb{R}^n$, $\det D^2_{xy} c \neq 0$.

The hypotheses (1.3)–(1.5) follow from the constructions of F, ϕ together with $f > 0, g > 0$. The mass balance condition (5.1) implies (1.8). The condition (A1) implies both (H1) and (H1*), and (A2) implies (H2).

Therefore, by Theorem 1.1 we have the existence of potentials (u, v) and optimal mapping T. The existence of potentials in optimal transportation was previously proved in [4, 5, 8]. By directly applying the formula (2.16), one obtains the optimal transportation equation

$$(5.4) \qquad \left| \det \left[D^2 u - D^2_{xx} c \right] \right| = |\det c_{x,y}| \frac{f}{g}.$$

We remark that in the linear case (5.2)–(5.3), both F in (1.1) and ϕ in (1.2) are linear in t, s variables, that is a border situation of F being concave and ϕ being convex in t, s, simultaneously.

There are numerous applicatons of the optimal transportation. Here we mention two important ones in geometric optics. In [25], Xu-Jia Wang showed that the far field reflector problem is an optimal transportation problem, and so is a linear optimization problem. The associated cost function $c(x, y) = -\log(1 - x \cdot y)$, where x, y are points on the unit sphere $\mathbb{S}^2$. Later on in [12] Gutiérrez and Huang showed that the far field refractor problem is also an optimal transporation problem. Let κ be the refractor index from the initial media to the target media. Then the associated cost function $c(x, y) = -\log(1 - \kappa x \cdot y)$ when $\kappa < 1$; and $c(x, y) = \log(\kappa x \cdot y - 1)$ when $\kappa > 1$, where x, y are points on the unit sphere $\mathbb{S}^n$. In the following subsections, we will consider more general reflector and refractor problems and show that they are in the class of the nonlinear optimization (1.1)–(1.2).

5.2. **Near field reflector problem with point source.** In [16] we showed that the near field reflector problem is a nonlienar optimization. For the convenience of the reader, we summaries the arguments as follows. Assume that the light emits from the origin O and passes through $\Omega \subset \mathbb{S}^n$ with a positive density $f \in L^1(\Omega)$. After being reflected from a surface Γ, the light will illuminate the target surface Ω^* in $\mathbb{R}^{n+1}$ with a prescribed positive density $g \in L^1(\Omega^*)$. Assume the energy conservation condition

$$(5.5) \qquad \int_\Omega f = \int_{\Omega^*} g.$$

Represent the reflector Γ in polar coordinate system as

$$\Gamma_\rho = \{X\rho(X) \,:\, X \in \Omega)\},$$

where ρ is a positive function. Recall that [15], Γ_ρ is *admissible* if at each point $X\rho(X) \in \Gamma$ there exists a *supporting ellipsoid*. Therefore, the radial function ρ satisfies

(5.6)
$$\rho(X) = \inf_{Y \in \Omega^*} \frac{p(Y)}{1 - \epsilon(p(Y))\langle X, \frac{Y}{|Y|}\rangle}, \quad X \in \Omega,$$

where p is the focal function on Ω^* and $\epsilon(p) = \sqrt{1 + p^2/|Y|^2} - p/|Y|$ is the eccentricity. Because there is an ellipsoid $E_{Y,p(Y)}$ supporting to Γ_ρ for each $Y \in \Omega^*$, we also have

(5.7)
$$p(Y) = \sup_{X \in \Omega} \rho(X)\left[1 - \epsilon(p(Y))\langle X, \frac{Y}{|Y|}\rangle\right], \quad Y \in \Omega^*.$$

Note that in (5.6) for each $X \in \Omega$ the infimum is achieved at some $Y \in \Omega^*$ and in (5.7) for each $Y \in \Omega^*$ the supremum is achieved at some $X \in \Omega$.

The relations (5.6)–(5.7) between the radial and focal functions of a reflector Γ are analogous to the classical relations between the radial and support functions for convex bodies, for example, see [21]. Inspired by that and [25], we set $\eta = 1/p$. Then the pair (ρ, η) satisfies the dual relation

(5.8)
$$\rho(X) = \inf_{Y \in \Omega^*} \frac{1}{\eta(Y)\left(1 - \epsilon(\eta(Y))\langle X, \frac{Y}{|Y|}\rangle\right)},$$

$$\eta(Y) = \inf_{X \in \Omega} \frac{1}{\rho(X)\left(1 - \epsilon(\eta(Y))\langle X, \frac{Y}{|Y|}\rangle\right)}.$$

Similarly to [25], we can now formulate the reflector problem to a nonlinear optimization (1.1)–(1.2) as follows. Let $u = \log \rho$ and $v = \log \eta$. Set the functional

(5.9)
$$I(u, v) = \int_{\Omega\times\Omega^*} F(X, Y, u, v)$$
$$= \int_{\Omega\times\Omega^*} f(X)u + g(Y)\left(v + \log(1 - \frac{\langle X, Y\rangle}{e^{-v} + \sqrt{|Y|^2 + e^{-2v}}})\right),$$

and the constraint set

$$K = \{(u, v) \in C(\Omega) \times C(\Omega^*) \,:\, \phi(X, Y, u, v) \le 0\},$$

with the constraint function

(5.10)
$$\phi(X, Y, u, v) = u + v + \log\left(1 - \frac{\langle X, Y\rangle}{e^{-v} + \sqrt{|Y|^2 + e^{-2v}}}\right).$$

We assume a further condition on domains Ω and Ω^*: Ω^* is contained in the cone $\mathcal{C}_V = \{tX \,:\, t > 0, X \in V\}$ for a domain $V \subset \mathbb{S}^n$ and

(5.11)
$$\overline{\Omega} \cap \overline{V} = \emptyset,$$

where $\overline{\Omega}$ and $\overline{V}$ denote the closure of Ω and V. It implies that there exists a small constant $\delta_0 > 0$, such that for any $X \in \Omega, Y \in \Omega^*$,

$$(5.12) \qquad\qquad -1 \le \langle X, \frac{Y}{|Y|} \rangle \le 1 - \delta_0.$$

Under the assumption (5.12), one can verify that (1.3)–(1.5) are satisfied by (5.9) and (5.10). The energy conservation condition (5.5) implies (1.8). To show the condition (H1), let $(u, v) \in K$ be a dual maximizing pair of I, (see Lemma 2.1). If there holds

$$\varphi_x(X, Y, v(Y)) = -Du(X), \quad \varphi(X, Y, v(Y)) = -u(X),$$

at $X \in \Omega, Y \in \Omega^*$, it was proved [16] that Y is the target point of light emitting along X, reflected at $Xe^{u(X)} \in \Gamma$ with unit normal

$$\gamma = \frac{(Du, 0) - (1 + Du \cdot x)X}{\sqrt{1 + |Du|^2 - (Du \cdot x)^2}}.$$

By the reflection law, $Y_r = X - 2\langle X, \gamma \rangle \gamma$,

$$Y = Xe^{u(X)} + Y_r|Y - Xe^{u(X)}|$$

is uniquely determined. It implies that $(\varphi_x, \varphi)(X, \cdot, \cdot)$ is one-to-one in $\Omega^* \times v(\Omega^*)$ for each $X \in \Omega$, see Remark 2.2.

As a consequent of Theorem 1.1, we have [16]

Corollary 5.1. *Assume that f, g satisfy (5.5). Suppose that Ω and Ω^* satisfy (5.11). Then there is a dual maximizing pair $(u, v) \in K$ satisfing*

$$I(u, v) = \sup_{(u,v) \in K} I(u, v),$$

where $I(u, v)$ is in (5.9), and the constraint ϕ is in (5.10). Moreover, $\rho = e^u$ is a solution of the reflector problem with given densities (Ω, f) and (Ω^*, g). (Note that the solutions need to be understood as generalized solutions.)

By directly applying the formula (2.16), we also obtain the PDE in the near field case [16], which was previously obtained by Karakhanyan and Wang in [15]. Assume that Ω^* is given implicitly by

$$(5.13) \qquad\qquad \Omega^* = \{Z \in \mathbb{R}^{n+1} : \psi(Z) = 0\}.$$

Suppose that Ω is a subset of upper unit sphere $\mathbb{S}^n_+ = \mathbb{S}^n \cap \{x_{n+1} > 0\}$. Let $X = (x, x_{n+1})$ be a parameterization of Ω, where $x_{n+1} = \sqrt{1 - |x|^2} =: \omega(x)$, and $x = (x_1, \cdots, x_n)$. For simplification, we define some auxiliary functions

$$(5.14) \qquad\qquad a = |D\rho|^2 - (\rho + D\rho \cdot x)^2,$$

$$(5.15) \qquad\qquad b = |D\rho|^2 + \rho^2 - (D\rho \cdot x)^2,$$

$$(5.16) \qquad\qquad t = \frac{\rho x_{n+1} - y_{n+1}}{\rho x_{n+1}}, \quad \beta = \frac{t}{(Y - X\rho) \cdot \nabla \psi},$$

and denote the matrix

$$(5.17) \qquad \mathcal{N} = \{\mathcal{N}_{ij}\}, \quad \mathcal{N}_{ij} = \delta_{ij} + \frac{x_i x_j}{1 - |x|^2}.$$

By computing in the local orthonormal frame, we obtain the equation as follows

Corollary 5.2. *The function ρ is a solution of*

$$(5.18) \qquad \left| det \left[D^2\rho - \frac{2}{\rho} D\rho \otimes D\rho - \frac{a(1-t)}{2t\rho}\mathcal{N} \right] \right| = \left| \frac{a^{n+1}}{t^n b\beta} \right| \frac{f}{2^n \rho^{2n+1}\omega^2 g|\nabla\psi|}.$$

Note that the matrix in equation (5.18) has a different sign to that in [15], since we calculate the absolute value of the determinant.

5.3. Near field reflector problem with parallel source.

The ideal reflection system can be described as follows: a parallel light emits from $\Omega \subset \mathbb{R}^n \times \{0\}$ along $e_{n+1} = (0, \cdots, 0, 1)$ with a positive density $f \in L^1(\Omega)$. After being reflected by the surface Γ in $\mathbb{R}^{n+1}$, it will illuminate the target domain $\Omega^* \subset \mathbb{R}^n \times \{0\}$ with the prescribed density $g \in L^1(\Omega^*)$. Assume the energy conservation condition

$$(5.19) \qquad \int_\Omega f = \int_{\Omega^*} g.$$

We represent the reflector Γ as graph $u|_\Omega$, namely

$$\Gamma_u = \{(x, u(x)) \, : \, x \in \Omega\},$$

where u is a positive function.

Consider the *"inverse" paraboloid* with focus at $y \in \Omega^*$ and the axial direction $-e_{n+1}$. The reflection property of these paraboloids is that: all the incident light from parallel source along the direction e_{n+1} will be reflected by the inverse paraboloids to the focus points y. Such an inverse paraboloid can be represented by $\Gamma_p = \{(x, p(x)) \, : \, x \in \mathbb{R}^n\}$ with

$$(5.20) \qquad p(x) = p_y(x) = \frac{1}{2v} - \frac{v}{2}|x - y|^2,$$

where $v > 0$ is a constant depending on y such that $p(x) > 0$ in Ω.

If we regard $v = v(y)$ as a function on Ω^*, we then have a family of paraboloids $\{\Gamma_{p_y} \, : \, y \in \Omega^*\}$. It is natural to construct the reflector Γ_u by the envelope of the family of paraboloids $\{\Gamma_{p_y} \, : \, y \in \Omega^*\}$.

In an ideal system, at each point $(x, u(x)) \in \Gamma_u$ on an *admissible* reflector Γ_u there exists a *supporting paraboloid*, namely for some $y \in \Omega^*$

$$u(x) = \frac{1}{2v(y)} - \frac{v(y)}{2}|x - y|^2, \quad \text{and} \quad u(x') \le \frac{1}{2v(y)} - \frac{v(y)}{2}|x' - y|^2 \quad \forall x' \in \Omega.$$

Hence, the defining function u and the dual function v satisfy the dual relation:

$$(5.21) \qquad u(x) = \inf_{y \in \Omega^*} \left\{ \frac{1}{2v(y)} - \frac{v(y)}{2}|x - y|^2 \right\}, \quad x \in \Omega,$$

$$(5.22) \qquad v(y) = \inf_{x \in \Omega} \left\{ \frac{-u + \sqrt{u^2 + |x - y|^2}}{|x - y|^2} \right\}, \quad y \in \Omega^*.$$

Note that in (5.21), if for $x \in \Omega$ the infimum is achieved at some $y = T(x) \in \Omega^*$, then in (5.22) at $y = T(x)$ the infimum is achieved at $x \in \Omega$. By (5.21) one has

$$Du(x) = -v(x - y),$$

and thus

$$|Du(x)|^2 = 1 - 2vu.$$

Since $u > 0, v > 0$, we have the natural restriction $|Du| < 1$ on Ω.

Based on (5.21)–(5.22) we can now formulate this problem to a nonlinear optimization problem (1.1)–(1.2). Set the functional

$$(5.23) \qquad \begin{aligned} I(u, v) &= \int_{\Omega \times \Omega^*} F(x, y, u, v) \\ &= \int_{\Omega \times \Omega^*} f(x)u(x) + g(y)\left(\frac{v}{2}|x - y|^2 - \frac{1}{2v} \right), \end{aligned}$$

and the constraint set

$$\mathcal{K} = \{(u, v) \in C(\Omega) \times C(\Omega^*) : \phi(x, y, u, v) \le 0\},$$

with the constraint function

$$(5.24) \qquad \phi(x, y, u, v) = u - \frac{1}{2v} + \frac{v}{2}|x - y|^2.$$

Note that ϕ can be written as (1.4) with $\varphi(x, y, v) = -\frac{1}{2v} + \frac{v}{2}|x - y|^2$. By differentiating, we have

$$(5.25) \qquad \begin{aligned} \varphi_x &= (x - y)v, \quad \varphi_y = -(x - y)v, \quad \varphi_{xx} = vI, \quad \varphi_{xy} = -vI, \\ \varphi_s &= \frac{1}{2v^2} + \frac{1}{2}|x - y|^2, \quad \varphi_{xs} = x - y, \quad etc. \end{aligned}$$

From Lemma 2.1 there exists a dual maximizing pair $(u, v) \in K$ of I. From (5.25) one can see that $(\varphi_x, \varphi)(x, \cdot, \cdot)$ is one-to-one in $\Omega^* \times v(\Omega^*)$ for each $x \in \Omega$, Remark 2.2. The energy conservation condition (5.19) implies (1.8). Alternatively, one can directly verify that the optimal mapping T associated to (u, v), determined by (2.6), is equal to the reflection mapping. This implies the condition (H1) due to the reflection law.

Proposition 5.1. *Let $(u, v) \in K$ be a dual maximizing pair of I, $|Du| < 1$. The associated optimal mapping T determined by (2.6) is equal to the reflection mapping T_r obtained by the reflection law.*

Proof. From (2.6) and (5.25), at $(x, y) = (x, T(x))$ we have

$$(5.26) \qquad\qquad x - y = -\frac{Du}{v}.$$

From (1.7), $u - \frac{1}{2v} + \frac{v}{2}|x - y|^2 = 0$, thus

$$(5.27) \qquad\qquad v = \frac{1 - |Du|^2}{2u}.$$

Note that $u > 0, v > 0$ and $|Du| < 1$. Combining (5.26) and (5.27), we obtain

$$(5.28) \qquad\qquad x - y = -\frac{2uDu}{1 - |Du|^2}.$$

This implies the optimal mapping T is given by

$$T(x) = x + \frac{2uDu}{1 - |Du|^2}.$$

Next, we calculate the reflection mapping T_r. At point $(x, u(x)) \in \Gamma_u$, from direct calculations the unit normal is

$$\gamma = \frac{(Du, -1)}{\sqrt{1 + |Du|^2}}.$$

By the reflection law, we have the reflected direction

$$(5.29) \qquad Y_r = e_{n+1} - 2\langle e_{n+1}, \gamma\rangle\gamma = \frac{(2Du, |Du|^2 - 1)}{|Du|^2 + 1}.$$

On the other hand, since the reflected ray meets the hyperplane $\mathbb{R}^n \times \{0\}$, we have

$$(5.30) \qquad\qquad Y_r = \frac{(y - x, -u)}{\sqrt{|y - x|^2 + u^2}}.$$

From (5.29)–(5.30) we obtain that

$$(5.31) \qquad y = T_r(x) = x + \frac{2uDu}{1 - |Du|^2} = T(x).$$

$\square$

Therefore, as a consequence of Theorem 1.1, we have

Corollary 5.3. *Assume that f, g satisfy the energy conservation condition (5.19). Then there is a dual maximizing pair $(u, v) \in K$ of I, where I, K are in (5.23), (5.24). Moreover, u is a (generalized) solution of the reflector problem with given densities (Ω, f) and (Ω^*, g).*

We now derive the equation for the potential function u by using the formula (2.16),

$$(5.32) \qquad\begin{aligned}\left|\det\left[D^2u + vI\right]\right| &= v^n\left|\det\left[I - \frac{1}{\varphi_s}(x - y)\otimes(x - y)\right]\right|\frac{f}{g}, \\ &= v^n\left|1 - \frac{1}{\varphi_s}|x - y|^2\right|\frac{f}{g},\end{aligned}$$

where in the second step we used the formula $\det\left[I + \xi \otimes \eta\right] = 1 + \xi \cdot \eta$ for any vectors $\xi, \eta \in \mathbb{R}^n$. Combining (5.25)–(5.28), we obtain that

$$(5.33) \qquad \phi_s = \frac{1}{2}\frac{|x-y|^2}{|Du|^2} + \frac{1}{2}|x-y|^2 = \frac{1+|Du|^2}{2|Du|^2}|x-y|^2.$$

Therefore, the equation (5.32) becomes

$$(5.34) \qquad \det\left[D^2u + \frac{1-|Du|^2}{2u}I\right] = \frac{(1-|Du|^2)^{n+1}}{(2u)^n(1+|Du|^2)}\frac{f}{g}.$$

The equation (5.34) was also obtained in [18] by directly computing the Jacobian of the reflection mapping. Denote the matrix

$$A(u, Du) = \frac{1-|Du|^2}{2u}I.$$

It satisfies the (A3) condition in [19] without the orthogonal restriction, provided that $u > 0$. The regularity of (5.34) follows from [18, 19]. Indeed the considerations in [19] stemmed from the treatment of the reflector antenna problem by Wang in [24], which can be represented as an optimal transportation problem on the sphere $\mathbb{S}^n$ with the cost function $c(x, y) = -\log(1 - x \cdot y)$.

5.4. **Near field refractor problem with point source.** The near field refractor problem has been studied by Gutiérrez and Huang [13]. Suppose the light emits from the origin surrounded by medium I with positive intensity $f(X)$ for $X \in \Omega$, where $\Omega \subset \mathbb{S}^n$. There is a surface $\mathcal{R}$, separates two homogeneous and isotropic media I and II, such that all rays refracted by $\mathcal{R}$ into medium II illuminate a target hypersurface Ω^* in $\mathbb{R}^{n+1}$ with positive intensity g on Ω^*. Assume that f, g satisfy the energy conservation condition

$$(5.35) \qquad \int_\Omega f = \int_{\Omega^*} g.$$

Let n_1, n_2 be the indices of refraction of media I, II, respectively, and $\kappa = n_2/n_1$. When $\kappa < 1$, the refracted rays tend to bent away from the normal, when $\kappa > 1$, the refracted rays tend to bent towards the normal.

There is a special interface surface $\mathcal{S}$ between media I and II, called Cartesian oval [13], that refracts all rays emitting from the origin O into the point Y. In the polar coordinates, represent

$$\mathcal{S} = \{X\rho_o(X) : X \in \mathbb{S}^n\}.$$

In the case $\kappa < 1$, by the Snell law of refraction one has

$$(5.36) \qquad \rho_o(X) = h(X, Y, p) = \frac{(p - \kappa^2 X \cdot Y) - \sqrt{\Delta(X \cdot Y)}}{1 - \kappa^2},$$

where p is the focal parameter, $\kappa|Y| < p < |Y|$ and $\Delta(t) = (p - \kappa^2 t)^2 - (1 - \kappa^2)(p^2 - \kappa^2|Y|^2)$ for $t \in \mathbb{R}$. For non-degenerate Cartesian ovals, there are some physical constraints for

refraction

$$(5.37) \qquad \kappa|Y| < p < |Y|, \quad \rho_o \le p, \quad \text{and } p \le X \cdot Y, \quad \forall X \in \Omega.$$

We refer the readers to [13] for more physical interpretations and detailed calculations.

If we regard $p = p(Y)$ as a focal function on Ω^*, we then have a family of Cartesian ovals. Represent the surface $\mathcal{R}$ in polar coordinate system as

$$\mathcal{R}_\rho = \{X\rho(X) : X \in \Omega\},$$

where ρ is a positive function. Recall that [13], $\mathcal{R}$ is a near field refractor if at each point $X\rho(X) \in \mathcal{R}$ there exists a supporting Cartesian oval, i.e., for some $Y \in \Omega^*$, $\rho(X') \le \rho_o(X', Y, p(Y))$ for all $X' \in \Omega$ with equality holds at $X' = X$. Therefore, the radial function ρ satisfies

$$(5.38) \qquad \rho(X) = \inf_{Y \in \Omega^*} \frac{(p - \kappa^2 X \cdot Y) - \sqrt{\Delta(X \cdot Y)}}{1 - \kappa^2}, \quad X \in \Omega.$$

From the energy conservation, for each $Y \in \Omega^*$ there is an oval $\rho_0(\cdot, Y, p(Y))$ supporting to $\mathcal{R}_\rho$. We also have

$$(5.39) \qquad p(Y) = \sup_{X \in \Omega} (1 - \kappa^2)\rho(X) + \kappa^2 X \cdot Y + \sqrt{\Delta(X \cdot Y)}, \quad Y \in \Omega^*.$$

The above relations are analogous to (5.6)–(5.7). By setting $\eta = 1/p$, the pair (ρ, η) satisfies the dual relation

$$(5.40) \qquad \rho(X) = \inf_{Y \in \Omega^*} \frac{1 - \eta(\kappa^2 X \cdot Y + \sqrt{\Delta(X \cdot Y)})}{\eta(1 - \kappa^2)},$$

$$(5.41) \qquad \eta(Y) = \inf_{X \in \Omega} \frac{1 - \eta(\kappa^2 X \cdot Y + \sqrt{\Delta(X \cdot Y)})}{\rho(1 - \kappa^2)}.$$

Similarly to (5.9) we now formulate the refractor problem to the following nonlinear optimization, which is more complicated than (5.9). Let $u = \log\rho$ and $v = \log\eta$. Set the functional

$$(5.42) \qquad \begin{aligned} I(u, v) &= \int_{\Omega \times \Omega^*} F(X, Y, u, v) \\ &= \int_{\Omega \times \Omega^*} f(X)u + g(Y)\left(v + \log\left(\frac{1 - \kappa^2}{1 - e^v(\kappa^2 X \cdot Y + \sqrt{\Delta(X \cdot Y)})}\right)\right), \end{aligned}$$

and the constraint set

$$K = \{(u, v) \in C(\Omega) \times C(\Omega^*) : \phi(X, Y, u, v) \le 0\},$$

with the constraint function

$$(5.43) \qquad \phi(X, Y, u, v) = u + v + \log\left(\frac{1 - \kappa^2}{1 - e^v(\kappa^2 X \cdot Y + \sqrt{\Delta(X \cdot Y)})}\right).$$

As in [13], we make the following assumptions on Ω and Ω^*, which are due to the physical constraints for refraction (5.37):

(R1) There exists τ with $0 < \tau < 1 - \kappa$ such that $X \cdot Y \geq (\kappa + \tau)|Y|$ for all $x \in \overline{\Omega}$ and all $Y \in \overline{\Omega}^*$.

(R2) Let $0 < r_0 \leq \frac{\tau}{1+\kappa}\mathrm{dist}(0, \overline{\Omega}^*)$ and consider the cone in $\mathbb{R}^{n+1}$

$$Q_{r_0} = \{tX \,:\, X \in \overline{\Omega}, 0 < t < r_0\}.$$

For each $\xi \in \mathbb{R}^n$ and for each $X \in Q_{r_0}$ we assume that $\overline{\Omega}^* \cap \{X + t\xi \,:\, t \geq 0\}$ contains at most one point. That is, for each $X \in Q_{r_0}$ each ray emanating from X intersects $\overline{\Omega}^*$ at most in one point.

Note that from (5.36)

(5.44) $\Delta(X \cdot Y) = \kappa^2 p^2 - 2\kappa^2(X \cdot Y)p + \kappa^2(X \cdot Y)^2 + \kappa^2(1 - \kappa^2)|Y|^2.$

Since $p \leq X \cdot Y$ by (5.37), $\Delta(X \cdot Y)$ is decreasing in p. Since $p = e^{-v}$, $\Delta(X \cdot Y)$ is increasing in v. Hence (5.42)–(5.43) satisfy the assumption (1.3) and (1.5), where the constant in (1.5) depends on κ, Ω, Ω^* and the assumptions (R1) and (R2). See Remark 4.2 in [13] for more physical interpretations of (R1) and (R2).

By the proof of Theorem 1.1 and Remark 2.1 we have the following existence result in the near field refractor problem, which was previously obtained in [13]. Note that the solutions need to be understood as generalized solutions.

Corollary 5.4. *Assume that f, g satisfy (5.35) and (R1)–(R2). Then given $P_0 \in Q_{r_0}$ with $0 < |P_0| \leq \left(\frac{1-\kappa}{1+\kappa}\right)^2 r_0$, there exists a dual maximizing pair $(u, v) \in K$ such that $u(P_0/|P_0|) = \log|P_0|$. Moreover, $\rho = e^u$ is a solution of the refractor problem such that $\mathcal{R}_\rho$ passes through the point P_0.*

Remark 5.1. In the case $\kappa > 1$, by the physical constraints for $|Y| < p < \kappa|Y|$ the refracting piece of Cartesian oval is given by

$$\mathcal{O}(Y, p) = \left\{ \rho_0(X, Y, p) \,:\, X \cdot Y \geq \frac{p + \sqrt{(\kappa^2 - 1)(\kappa^2|Y|^2 - p^2)}}{\kappa^2} \right\}$$

with

(5.45) $\rho_0(X, Y, p) = \dfrac{(\kappa^2 X \cdot Y - p) - \sqrt{(\kappa^2 X \cdot Y - b)^2 - (\kappa^2 - 1)(\kappa^2|Y|^2 - p^2)}}{\kappa^2 - 1}.$

We have similar existence results by replacing the assumptions (R1), (R2) by the following ones:

(R3) $\inf_{X \in \overline{\Omega}, Y \in \overline{\Omega}^*} X \cdot \frac{Y}{|Y|} \geq \frac{1}{\kappa} + \tau$, for some $0 < \tau < 1 - \frac{1}{\kappa}$.

(R4) Let $0 < r_0 < \frac{\kappa^2 \tau^2}{4(\kappa-1)^2} \inf_{Y \in \Omega^*} |Y|$ and consider the cone in $\mathbb{R}^{n+1}$

$$Q_{r_0} = \{tX \,:\, X \in \overline{\Omega}, 0 < t < r_0\}.$$

For each $\xi \in \mathbb{R}^n$ and for each $X \in Q_{r_0}$ we assume that $\overline{\Omega}^* \cap \{X + t\xi : t \geq 0\}$ contains at most one point. That is, for each $X \in Q_{r_0}$ each ray emanating from X intersects $\overline{\Omega}^*$ at most in one point.

5.5. Near field refractor problem with parallel source.

Recently, Gutiérrez and Tournier studied the parallel refractor problem [14], which can be described as follows. Suppose that a parallel light emits from $\Omega \subset \mathbb{R}^n \times \{0\}$ along $e_{n+1} = (0, \cdots, 0, 1)$ with positive intensity $f \in L^1(\Omega)$, Ω^* is a hypersurface in $\mathbb{R}^{n+1}$, which is referred to as the target domain. Suppose that Ω and Ω^* are surrounded by two homogeneous and isotropic media I and II, respectively. One seeks an optical surface $\mathcal{R}$ interface between media I and II, such that all rays refracted by $\mathcal{R}$ into medium II are received at the surface Ω^*, and the prescribed radiation intensity received at each point $Y \in \Omega^*$ is $g(Y)$. Assume the energy conservation condition

$$(5.46) \qquad \int_\Omega f = \int_{\Omega^*} g.$$

Let n_1, n_2 be the indices of refraction of media I, II, respectively, and $\kappa = n_1/n_2$. We assume that media II is denser than media I, that is, $\kappa < 1$. The case when $\kappa > 1$ can be treated in a similar way but the geometry of surface changes [12, 14]. For simplicity, we assume that $\Omega^* \subset \{y_{n+1} = h\}$ for a constant $h > 0$, and denote $Y = (y, h)$ for points on Ω^* and $X = (x, 0)$ for points on Ω.

Consider the lower part of *"inverse" ellipsoid of revolution* with focus at $y \in \Omega^*$ and the axial direction $-e_{n+1}$. It has the uniform refracting property, namely all rays from the parallel source Ω along e_{n+1} will be refracted to the focus points y. Explicitly, it is the graph of the function [14]

$$(5.47) \qquad \rho_{y,v}(x) = h - \frac{\kappa v}{1 - \kappa^2} - \sqrt{\frac{v^2}{(1 - \kappa^2)^2} - \frac{|x - y|^2}{1 - \kappa^2}},$$

where v is a constant satisfying $\frac{h(1 - \kappa^2)}{\kappa} \leq v \leq \frac{h(1 - \kappa^2)(1 + \kappa)^2}{\kappa^3}$. The function $\rho_{y,v}$ is defined on the ball $B_{v/\sqrt{1 - \kappa^2}}(y)$.

If we regard $v = v(y)$ as a function on Ω^*, we then have a family of "inverse" ellipsoids. Represent the refractor Γ as graph $u|_\Omega$ for $u > 0$, namely

$$\Gamma_u = \{(x, u(x)) : x \in \Omega\}.$$

In an ideal system, at each point $(x, u(x)) \in \Gamma_u$ on an *admissible* refractor Γ_u there exists a *supporting ellipsoid*, i.e., for some $y \in \Omega^*$

$$u(x) = h - \frac{\kappa v(y)}{1 - \kappa^2} - \sqrt{\frac{v(y)^2}{(1 - \kappa^2)^2} - \frac{|x - y|^2}{1 - \kappa^2}},$$

$$u(x') \leq h - \frac{\kappa v(y)}{1 - \kappa^2} - \sqrt{\frac{v(y)^2}{(1 - \kappa^2)^2} - \frac{|x' - y|^2}{1 - \kappa^2}} \quad \forall x' \in \Omega.$$

Similarly we can formulate this problem to a nonlinear optimization problem (1.1)–(1.2). Set the functional

(5.48)
$$I(u, v) = \int_{\Omega \times \Omega^*} F(x, y, u, v)$$
$$= \int_{\Omega \times \Omega^*} f(x) u + g(y) \left(\frac{\kappa v(y)}{1 - \kappa^2} + \sqrt{\frac{v(y)^2}{(1 - \kappa^2)^2} - \frac{|x - y|^2}{1 - \kappa^2}} \right),$$

and the constraint set

$$K = \{(u, v) \in C(\Omega) \times C(\Omega^*) \; : \; \phi(X, Y, u, v) \leq 0\},$$

with the constraint function

(5.49)
$$\phi(X, Y, u, v) = u + \frac{\kappa v(y)}{1 - \kappa^2} + \sqrt{\frac{v(y)^2}{(1 - \kappa^2)^2} - \frac{|x - y|^2}{1 - \kappa^2}} - h.$$

As in [14] we need the following assumptions on the relative position of Ω and Ω^*, which are due to the physical constraints for refraction:

(A) There exists $0 < \delta < 1$ such that $\Omega \subset B_{\delta h \sqrt{1 - \kappa^2}/\kappa}(y)$ for all $y \in \Omega^*$.

(B) Set $M = h((1 + \kappa)^3/\kappa^3 - 1)$. Assume that for all $x \in \Omega \times [-M, 0]$ and for all $\gamma \in \mathbb{S}^n$, the ray $\{x + t\gamma : t > 0\}$ intersects Ω^* in at most one point.

The first condition is equivalent to the assumption that there exists $0 < \beta < 1$ such that $\langle -e_{n+1}, \frac{X - Y}{|X - Y|} \rangle \geq \beta$ for all $Y \in \Omega^*$ and $x \in \Omega$ [14].

As in the previous example one can verify the hypotheses of Theorem 1.1. Therefore, by the proof of Theorem 1.1 and Remark 2.1 we have the existence result:

Corollary 5.5. *Assume that f, g satisfy (5.46) and (A)–(B). Then for $x_0 \in \Omega$ and $t \leq -\beta$ there exists a parallel refractor u satisfying $u(x_0) = t$.*

Note that the solutions need to be understood as generalized solutions as before. This existence result was previously obtained in [14], where they first consider the discrete case when the target is a set of points, then use an approximation to obtain the existence in the general case.

Remark 5.2. In addition to the examples arising in reflectors and refractors, there are many other nonlinear optimization problems with potentials. For example, one can perturb the linear optimization problem, such as the optimal transportation, to get a nonlinear one. Moreover, similarly to [20] one can show that the objective functional of any solvable linear optimization problem can be perturbed by a differentiable, convex or Lipschitz continuous nonlinear functional in such a way that (i) a solution of the original linear problem is a local or global solution of the perturbed nonlinear problem; (ii) each global solution of the perturbed nonlinear problem is also a solution of the linear problem.

REFERENCES

[1] Ambrosio, L., Lecture notes on optimal transport problems, *Mathematical aspects of evolving interfaces (Funchal, 2000),* 1–52, Lecture Notes in Math., 1812, *Springer, Berlin,* 2003.

[2] Bakelman, I., *Convex analysis and nonlinear geometric elliptic equations.* Springer, Berlin, Heidelberg, 1994.

[3] Bertsekas, D. P., *Nonlinear programming*, 2nd edition. Athena Scientific, 1999.

[4] Brenier, Y., Polar factorization and monotone rearrangement of vector-valued functions, *Comm. Pure Appl. Math.* 44 (1991), 375–417.

[5] Caffarelli, L., Allocation maps with general cost functions, in *Partial Differential Equations and Applications* (P. Marcellini, G. Talenti, and E. Vesinitini eds). Lecture Notes in Pure and Appl. Math. 177 (1996), 29–35.

[6] Cheng, S. Y. and Yau, S. T., On the regularity of the Monge-Ampère equation $\det(\partial^2 u/\partial x_i \partial x_j) = F(x, u)$, *Comm. Pure Appl. math.* 30 (1977), 41–68.

[7] Evans, L. C., Partial differential equations and Monge-Kantorovich mass transfer, in *Current developments in mathematics*, Int. Press, 1997 pp. 65–126.

[8] Gangbo, W. and McCann, R. J., Optimal maps in Monge's transport problem, *C. R. Acad. Sci. Paris Sér. I. Math.* 321 (1995), 1653–1658.

[9] Gilbarg, D. and Trudinger, N., *Elliptic partial differential equations of second order.* Springer-Verlag, Berlin, 1983.

[10] Goh, C. J. and Yang, X. Q., *Duality in optimization and variational inequalities*, Optimization Theory and Applications, 2. *Taylor & Francis, Ltd., London,* 2002.

[11] Gutiérrez, C., *The Monge-Ampère equation.* Progress in Nonlinear Differential Equations and their Applications, 44. Birkhäuser Boston, Inc., Boston, MA, 2001.

[12] Gutiérrez, C. E. and Huang, Q., The refractor problem in reshaping light beams, *Arch. Rational Mech. Anal.*, 193 (2009), 423–443.

[13] Gutiérrez, C. E. and Huang, Q., The near field refractor, preprint.

[14] Gutiérrez, C. E. and Tournier, F., The parallel refractor, *Developments in Mathematics*, to appear.

[15] Karakhanyan, A. and Wang, X.-J., On the reflector shape design, *J. Diff. Geom.*, 84 (2010), 561–610.

[16] Liu, J., Light reflection is nonlinear optimization, *Calc. Var. and PDEs*, to appear.

[17] Liu, J., A class of nonlinear optimization problems with potentials II, in preparation.

[18] Liu, J. and Trudinger, N. S., On the classical solvability of near field reflector problems, preprint.

[19] Ma, X. N.; Trudinger, N. S. and Wang, X.-J., Regularity of potential functions of the optimal transportation problem, *Arch. Rat. Mech. Anal.*, 177 (2005), 151–183.

[20] Mangasarian, O. L. and Meyer, R. R., Nonlinear perturbation of linear programs, *SIAM J. Control Optim.* 17 (1979), 745–752.

[21] Schneider, R., *Convex Bodies.* The Brunn-Minkowski Theory. *Cambridge University Press*, Cambridge, 1993.

[22] Urbas, J., *Mass transfer problems*, Lecture Notes, Univ. of Bonn, 1998.

[23] Villani, C., *Optimal transport. Old and new.* Grundlehren Math. Wiss., Vol. 338, *Springer-Verlag, Berlin*, 2009.

[24] Wang, X.-J., On the design of a reflector antenna, *Inverse problems* 12 (1996), 351–375.

[25] Wang, X.-J., On the design of a reflector antenna II, *Calc. Var. and PDEs* 20 (2004), 329–341.

Department of Mathematics, Princeton University, Fine Hall, Washington Road, Princeton, NJ 08544-1000, USA.

E-mail address: jiakunl@math.princeton.edu